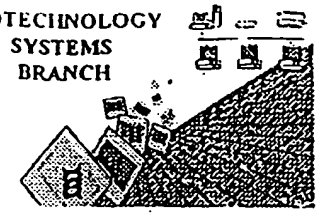
BIOTECHNOLOGY SYSTEMS BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/796,669
Source: IFW0
Date Processed by STIC: 3/18/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221
Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS
http://www.uspto.gov/web/offices/pac/checker/chkr41note.htm

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses

1. EFS-Bio (<http://www.uspto.gov/ebc/efs/downloads/documents.htm> , EFS Submission User Manual - ePAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10/796,669

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1 ✓	Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2	Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3	Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4 ✓	Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5	Variable Length	Sequence(s) ____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6	PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7	Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8	Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9	Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10	Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11	Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 07/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12	PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13	Misuse of n/Xaa	"n" can only represent a single nucleotide; "Xaa" can only represent a single amino acid

IFWO

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

```
      3 <110> APPLICANT: Jin-Town Wang
      4       Tzu-Lung Lin
W--> 5 <120> TITLE OF INVENTION: A TYPE II RESTRICTION ENDONUCLEASE AND APPLICATION THEREOF
C--> 6 <140> CURRENT APPLICATION NUMBER: US/10/796,669
C--> 6 <141> CURRENT FILING DATE: 2004-03-09
W--> 0 <130> FILE REFERENCE:
  > 6 <160> NUMBER OF SEQ ID: 5
      7 <170> SOFTWARE: MICROSOFT WORD 2000

ERRORED SEQUENCES

E--> 8 <210> SEQ ID NO: SEQ ID NO:1
     9 <211> LENGTH: LENGTH:5
     10 <212> TYPE: DNA
     11 <213> ORGANISM: ORGANISM: Helicobacter pylori
E--> 13 <400> SEQUENCE: SEQUENCE:1
E--> 14 ccatc      5
E--> 16 <210> SEQ ID NO: SEQ ID NO:2
     17 <211> LENGTH: LENGTH:1617
     18 <212> TYPE: DNA
     19 <213> ORGANISM: ORGANISM: Helicobacter pylori
     21 <300> PUBLICATION INFORMATION:
     22 <308> DATABASE ACCESSION NO: DDBJ/EMBL/Genbank;Accession No.:AB118944
  > 24 <300> PUBLICATION INFORMATION: SEQUENCE:2
  > 25 atg act aaa aaa ccg gca cga aaa att tta agc ttt tca
     26  acc acc atg cga aac cct aaa  60
E--> 28 aga ata gga caa ttt tta gct gtt tta gga aag ttt gaa
     29  aat caa atc ctt aaa tct tca  120
E--> 31 ata atc atg caa att atc aaa tcc gtt ttg gct cat agg
     32  ctt tat aga cct act tct ctc  180
E--> 34 aat caa aat aaa gaa ttg aaa gaa aaa ttt gac tcc aat
     35  gaa tat gtc ttt agc gat gaa  240
E--> 37 gag tta gaa cgc att ata gaa ata tcc cca caa aat cat
     38  aaa gaa atg ggc ttt gag cat  300
E--> 40 gga tgg gaa agc cgg ttt gac act tgg tat aag ctt atg
     41  tgt gag ttt ggt ttt tgc tac  360
E--> 43 tat gca aaa tat gag aaa ata ctc atc agc gat agc gct
     44  aag atg ctt att ctt gct tat  420
E--> 46 tac aat aaa gaa aac gat gct ttt aaa gaa agc gtt gat
     47  gaa agc gta gtt ggg gct ata  480
E--> 49 ttt tta aac gct ctg tct aaa tat gaa gta gga aac cct
     50  tac aaa aag aat tta aac cat  540
```

delete, please do not use alpha numeric headings in a sequence.

Does Not Comply
Corrected Diskette Needed
(pg. 1-7)

please insert

delete, same error

delete

delete, same error

(please see Attached sample Sequence Listing)

IF you include <308> With a response, then you must include <309> with a response.

please group the Nucleotides in sets of 10, for example

Please see item #1 on error summarysheet.

Please see item #4 on error summary sheet.

ttatagtagt tgtaaggatt

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

```
E--> 52 aac aac cct ttc aaa cta ttg ctc tcg ctt tta aaa cga
     53  ctc aaa aat gcc cat cta acc  600
E--> 55 ccc cta tct gtc aaa gaa atc cct att tta ctt tgt tgg
     56  aaa gac gat aac gct aat ggg  660
E--> 58 ctt tat gac tac att att cgt tta aga caa gaa atc gtt
     59  act atc aat aaa aca gaa ttc  720
E--> 61 agc tac tca gat gaa ttt atc tat gaa aaa tgc cta aaa
     62  ctt tta gaa agt gtt aat aaa  780
E--> 64 aca cga ttt aaa atg agc caa atc act aac gaa gcc gtt
     65  gat gaa tac att aga aaa atg  840
E--> 67 cgt att aca gga ctt att tca ttg cgt ggt aat ggt agg
     68  ttt att gat att aat act aat  900
E--> 70 gaa aat aat aaa ata gat tac att tta caa acc cat aag
     71  gct ttt aaa ggg gat tat tta  960
E--> 73 aac gac act caa gct aac aaa ctc gcc ttt ttt aac tac
     74  atg gcg atc gtg gat agc ttt  1020
E--> 76 ctt gtt agt gtt act cca atc agc gct aat gag agc gtt
     77  aaa tca agc aaa ttg aat gaa  1080
E--> 79 cta gca aac act tat act aaa gat ttt atc aag caa gaa
     80  tta ctc att act tgt aac aag  1140
E--> 82 caa gaa tca aaa gat agt ttt tta aga ctc att gat aaa
     83  cct tta cgc tta gaa ttt tta  1200
E--> 85 agc gct att ttc ttg aaa caa cat ttt gaa aat tta agc
     86  gtg ata ccc aat tat aaa agc  1260
E--> 88 gat gat gaa ggc ttg ccc gta tac aca gca agc ggt aat
     89  aaa cct gat att gta gct atg  1320
E--> 91 gac aca aaa gcc caa agt tat ata gaa gtg agc ttg att
     92  aga gac aga agt caa agt acc  1380
E--> 94 ttg gaa atg ata cct att gcc aga cat tta aaa gaa ttg
     95  att aaa aat agc acc gat att  1440
E--> 97 aga gaa aaa ttt agt gtt ttt gta gct cca aat atc cat
     98  gat gat gcc aaa gaa tat gcg  1500
E--> 100 gaa ttt gcc caa ttc aaa gac aat att aat ata tgt tgt
     101  tat gct att aat gat ttt atc  1560
  > 103 aaa aaa gta gaa aac agc ata gaa tgg tta cag atc aat
E--> 104  gac cat ttg aaa gct taa  1617
E--> 107 <210> SEQ ID NO: SEQ ID NO:3 delete,
     108 <211> LENGTH: LENGTH:538
     109 <212> TYPE: PRT delete
     110 <213> ORGANISM: ORGANISM: Helicobacter pylori
E--> 113 <400> SEQUENCE: SEQUENCE:3 delete
     114 Met Thr Lys Lys Pro Ala Arg Lys Ile Leu Ser Phe Ser
E--> 115  Thr Thr Met
E--> 116 1    5     10
E--> 117  15
     119 Arg Asn Pro Lys Arg Ile Gly Gln Phe Leu Ala Val Leu
E--> 120  Gly Lys Phe
E--> 121 20    25    30
     123 Glu Asn Gln Ile Leu Lys Ser Ser Ile Ile Met Gln Ile
```

please group Nucleotide in sets of 10.

See item #1 on error summary sheet.

Do not use alpha numeric headings.

please see item #1 on error summary sheet.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

E--> 124 Ile Lys Ser
E--> 125 35 40 45
127 Val Leu Ala His Arg Leu Tyr Arg Pro Thr Ser Leu Asn
E--> 128 Gln Asn Lys
E--> 129 50 55 60
131 Glu Leu Lys Glu Lys Phe Asp Ser Asn Glu Tyr Val Phe
E--> 132 Ser Asp Glu
E--> 133 65 70 75
E--> 134 80
136 Glu Leu Glu Arg Ile Ile Glu Ile Ser Pro Gln Asn His
E--> 137 Lys Glu Met
E--> 138 85 90 95
140 Gly Phe Glu His Gly Trp Glu Ser Arg Phe Asp Thr Trp
E--> 141 Tyr Lys Leu
E--> 142 100 105 110
144 Met Cys Glu Phe Gly Phe Cys Tyr Tyr Ala Lys Tyr Glu
E--> 145 Lys Ile Leu
E--> 146 115 120 125
148 Ile Ser Asp Ser Ala Lys Met Leu Ile Leu Ala Tyr Tyr
E--> 149 Asn Lys Glu
E--> 150 130 135 140
152 Asn Asp Ala Phe Lys Glu Ser Val Asp Glu Ser Val Val
E--> 153 Gly Ala Ile
E--> 154 145 150 155
E--> 155 160
157 Phe Leu Asn Ala Leu Ser Lys Tyr Glu Val Gly Asn Pro
E--> 158 Tyr Lys Lys
E--> 159 165 170 175
161 Asn Leu Asn His Asn Asn Pro Phe Lys Leu Leu Leu Ser
E--> 162 Leu Leu Lys
E--> 163 180 185 190
165 Arg Leu Lys Asn Ala His Leu Thr Pro Leu Ser Val Lys
E--> 166 Glu Ile Pro
E--> 167 195 200 205
169 Ile Leu Leu Cys Trp Lys Asp Asp Asn Ala Asn Gly Leu
E--> 170 Tyr Asp Tyr
E--> 171 210 215 220
173 Ile Ile Arg Leu Arg Gln Glu Ile Val Thr Ile Asn Lys
E--> 174 Thr Glu Phe
E--> 175 225 230 235
E--> 176 240
178 Ser Tyr Ser Asp Glu Phe Ile Tyr Glu Lys Cys Leu Lys
E--> 179 Leu Leu Glu
E--> 180 245 250 255
182 Ser Val Asn Lys Thr Arg Phe Lys Met Ser Gln Ile Thr
E--> 183 Asn Glu Ala
E--> 184 260 265 270
186 Val Asp Glu Tyr Ile Arg Lys Met Arg Ile Thr Gly Leu
E--> 187 Ile Ser Leu

please see item #1 on error summary sheet.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

E--> 188 275 280 285
190 Arg Gly Asn Gly Arg Phe Ile Asp Ile Asn Thr Asn Glu
E--> 191 Asn Asn Lys
E--> 192 290 295 300
194 Ile Asp Tyr Ile Leu Gln Thr His Lys Ala Phe Lys Gly
E--> 195 Asp Tyr Leu
E--> 196 305 310 315
E--> 197 320
200 Asn Asp Thr Gln Ala Asn Lys Leu Ala Phe Phe Asn Tyr
E--> 201 Met Ala Ile
E--> 202 325 330 335
204 Val Asp Ser Phe Leu Val Ser Val Thr Pro Ile Ser Ala
E--> 205 Asn Glu Ser
E--> 206 340 345 350
208 Val Lys Ser Ser Lys Leu Asn Glu Leu Ala Asn Thr Tyr
> 209 Thr Lys Asp
E--> 210 355 360 365
212 Phe Ile Lys Gln Glu Leu Leu Ile Thr Cys Asn Lys Gln
E--> 213 Glu Ser Lys
E--> 214 370 375 380
216 Asp Ser Phe Leu Arg Leu Ile Asp Lys Pro Leu Arg Leu
E--> 217 Glu Phe Leu
E--> 218 385 390 395
E--> 219 400
221 Ser Ala Ile Phe Leu Lys Gln His Phe Glu Asn Leu Ser
E--> 222 Val Ile Pro
E--> 223 405 410 415
225 Asn Tyr Lys Ser Asp Asp Glu Gly Leu Pro Val Tyr Thr
E--> 226 Ala Ser Gly
E--> 227 420 425 430
229 Asn Lys Pro Asp Ile Val Ala Met Asp Thr Lys Ala Gln
E--> 230 Ser Tyr Ile
E--> 231 435 440 445
233 Glu Val Ser Leu Ile Arg Asp Arg Ser Gln Ser Thr Leu
> 234 Glu Met Ile
E--> 235 450 455 460
237 Pro Ile Ala Arg His Leu Lys Glu Leu Ile Lys Asn Ser
E--> 238 Thr Asp Ile
E--> 239 465 470 475
E--> 240 480
243 Arg Glu Lys Phe Ser Val Phe Val Ala Pro Asn Ile His
E--> 244 Asp Asp Ala
E--> 245 485 490 495
247 Lys Glu Tyr Ala Glu Phe Ala Gln Phe Lys Asp Asn Ile
E--> 248 Asn Ile Cys
E--> 249 500 505 510
251 Cys Tyr Ala Ile Asn Asp Phe Ile Lys Lys Val Glu Asn
E--> 252 Ser Ile Glu
E--> 253 515 520 525

- Same errors, see item #1 on error summary sheet.

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

```
     255 Trp Leu Gln Ile Asn Asp His Leu Lys Ala
E--> 256 530     535
     259 <210> SEQ ID NO: 4
     260 <211> LENGTH: 780
     261 <212> TYPE: DNA
     262 <213> ORGANISM: ORGANISM: Helicobacter pylori
     264 <300> PUBLICATION INFORMATION:
     265 <308> DATABASE ACCESSION NO: DDBJ/EMBL/Genbank;Accession No.:AB118944
W--> 267 <300> PUBLICATION INFORMATION: 4
E--> 268 atg ggg caa gac gct gat ttt aaa gcg ctt gaa gaa ctg
     269  aaa gaa tac ttt aat caa gct  60
E--> 271 tta aag cta gaa gaa aat tat ttt agc caa cat ttt agc
     272  aac aag ttt ttc agc tat aaa  120
E--> 274 gat tgt gtc aaa atc ggt agc att aga gag cat ata gaa
     275  agc tta aac tta gat aaa tta  180
  > 277 aat aaa gat att tta tta aca agc ctg att tat tca atg
     278  gat aag ata gct aac acg gta  240
E--> 280 ggg cat tat gaa gct tat agg aaa aaa gag att ttg caa
     281  gat aga ttt att ttt gag ctt  300
E--> 283 att agc cct ata aaa cat gat aaa aat atc atg ata gag
     284  aga aaa aac gct aac gaa ttg  360
E--> 286 gct aaa acc tta aaa ata gac tta gtc ttt att gat cct
     287  cca tac aat tca agg caa tac  420
E--> 289 agc cgg ttt tat cat ctc tat gaa aac cta gtg cag tgg
     290  aaa aaa ccc aaa ctc tat gga  480
E--> 292 aca gct tta aag cca tca tgc gag aac atg agc gaa tat
     293  tgc cgc tct aat gcc aag aaa  540
E--> 295 gaa ttg agc gat tta att gaa aaa cta gat tgt aaa agg
     296  att gct tta act tat aat aat  600
E--> 298 acc tat aac tct aag tct agc tct tcg caa aat aaa ata
     299  ggc ttt aaa gat tta gtg gaa  660
E--> 301 att ttg agt caa aaa gga aaa tta agc gtt aaa gaa aag
     302  gct cat agt ttt ttt aat tca  720
E--> 304 gga aaa act gat ttt aaa gag cat aaa gaa ttt tta ttt
     305  ata gtg gaa gtg aaa cct tga  780
E--> 308 <210> SEQ ID NO: SEQ ID NO:5
     309 <211> LENGTH: LENGTH 846
     310 <212> TYPE: DNA
     311 <213> ORGANISM: ORGANISM Helicobacter pylori
     313 <300> PUBLICATION INFORMATION:
     314 <308> DATABASE ACCESSION NO: DDBJ/EMBL/Genbank;Accession No.:AB118944
W--> 316 <300> PUBLICATION INFORMATION: SEQUENCE:5
E--> 317 atg cca caa ctc aat aag cta ttc cca aat aac att aat
     318  caa ttt att gag cct ttt gtg  60
E--> 320 ggt ggg ggt agc gtg ttt tta aac act aag gct aag aga
     321  tac tta gct aat gac ata gat  120
E--> 323 act aat att atc aat tta cat aaa act tta agc aag ttc
     324  aat gtt tgt gag ctt ttt gat  180
E--> 326 gaa ttg tct aaa att atc att cat tat ggc ttg tct ttc
```

same error

delete, same error

IF you include <308> then you must include <309> with response.

same errors

delete, same errors

same errors, please insert also <309>

same errors

RAW SEQUENCE LISTING
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:36

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

```
     327  tct ttt aag ggg att atg gcc  240
E--> 329 cct gat gaa tta aaa aaa caa tat ata aaa act tac tac
     330  gcc aaa tac aat aaa ata gct  300
E--> 332 tat gaa aaa cta agg gct gat ttt aac tcc aat caa aac
     333  aac atg ctt tat ttg tat ttg  360
E--> 335 ctt tta att tat gga ttt aat cac atg att aga ttt aat
     336  tct aaa ggg ctt ttt aat tta  420
E--> 338 cct gtg ggt aat gtg gat ttc aat gaa aat gtt tat aat
     339  gcc cta aaa aac tac ata gat  480
E--> 341 ttt ata cag caa aac acc att att ttt cac aat gat gat
     342  tat att gat ttt ctt aac cac  540
E--> 344 acc act tat tta aaa gat gat tat gtt tat ttt gac ccc
     345  cct tat tta atc tcc aat agt  600
E--> 347 gaa tac aac aag tta tgg gat agc gat aat gag ata gcc
     348  tta tat ggt gtt tta gat agc  660
   > 350 cta gat aaa aag gga gtt tta ttt ggt ata act aat ctt
     351  att tat cac aag gga gag act  720
E--> 353 aat ttt att tta aaa gaa tgg gct aaa aaa tat tat att
     354  ttt aat atc aaa agt aat tat  780
E--> 356 atc agt tat aat gac aat act att aaa gaa gat agt caa
     357  gaa atc ttt gta act aat tat  840
     359 agg tga  846
E--> 3621
```

some errors

Delete

VERIFICATION SUMMARY
PATENT APPLICATION: US/10/796,669

DATE: 03/18/2004
TIME: 15:02:37

Input Set : A:\PTO.DA.txt
Output Set: N:\CRF4\03172004\J796669.raw

L:5 M:283 W: Missing Blank Line separator, <120> field identifier
L:6 M:270 C: Current Application Number differs, Replaced Current Application No
L:6 M:271 C: Current Filing Date differs, Replaced Current Filing Date
L:0 M:201 W: Mandatory field data missing, <130> FILE REFERENCE
L:6 M:283 W: Missing Blank Line separator, <160> field identifier
L:8 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:8 M:283 W: Missing Blank Line separator, <210> field identifier
L:13 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:14 M:252 E: No. of Seq. differs, <211> LENGTH:Input:0 Found:5 SEQ:0
L:16 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:24 M:256 W: Invalid Numeric Header Field, Identifier <309> Expected, SEQ:0
L:24 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:25 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:39 SEQ:0
M:254 Repeated in SeqNo=0
L:104 M:252 E: No. of Seq. differs, <211> LENGTH:Input:0 Found:1617 SEQ:0
L:107 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:113 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:115 M:332 E: (32) Invalid/Missing Amino Acid Numbering, SEQ ID:0
M:332 Repeated in SeqNo=0
L:256 M:252 E: No. of Seq. differs, <211> LENGTH:Input:0 Found:538 SEQ:0
L:267 M:256 W: Invalid Numeric Header Field, Identifier <309> Expected, SEQ:4
L:268 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:39 SEQ:4
M:254 Repeated in SeqNo=4
L:308 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQ ID NO
L:316 M:256 W: Invalid Numeric Header Field, Identifier <309> Expected, SEQ:0
L:316 M:212 E: (34) Invalid or duplicate Sequence ID Number, SEQUENCE ID NOS:0 differs:4
L:317 M:254 E: No. of Bases conflict, LENGTH:Input:0 Counted:39 SEQ:0
M:254 Repeated in SeqNo=0
L:362 M:252 E: No. of Seq. differs, <211> LENGTH:Input:0 Found:846 SEQ:0

(Sample Sequence Listing)

<110> Smith, John; Smithgene Inc.

<120> Example of a Sequence Listing

<130> 01-00001

<140> PCT/EP98/00001
<141> 1998-12-31

<150> US 08/999,999
<151> 1997-10-15

<160> 4

<170> PatentIn version 2.0

<210> 1
<211> 389
<212> DNA
<213> Paramecium sp.

<220>
<221> CDS
<222> (279)...(389)

<300>
<301> Doe, Richard
<302> Isolation and Characterization of a Gene Encoding a Protease from Paramecium sp.
<303> Journal of Genes
<304> 1
<305> 4
<306> 1-7
<307> 1988-06-31
<308> 123456
<309> 1988-06-31

<400> 1

agctgtagtc attcctgtgt cctcttctct ctgggcttct caccctgcta atcagatctc 60

agggagagtg tcttgaccct cctctgcctt tgcagcttca caggcaggca ggcaggcagc 120

tgatgtggca attgctggca gtgccacagg cttttcagcc aggcttaggg tgggttccgc 180

cgcggcgcgg cggccctct cgcgctcctc tcgcgcctct ctctcgctct cctctcgctc 240

10/796,664

(Sample Sequence Listing)

```
ggacctgatt  aggtgagcag  gaggaggggg  cagttagc      atg  gtt  tca  atg  ttc  agc   296
                                                  Met  Val  Ser  Met  Phe  Ser
                                                   1                   5

ttg  tct  ttc  aaa  tgg  cct  gga  ttt  tgt  ttg  ttt  gtt  tgt  ttg  ttc  caa   344
Leu  Ser  Phe  Lys  Trp  Pro  Gly  Phe  Cys  Leu  Phe  Val  Cys  Leu  Phe  Gln
                10                       15                  20

tgt  ccc  aaa  gtc  ctc  ccc  tgt  cac  tca  tca  ctg  cag  ccg  aat  ctt        389
Cys  Pro  Lys  Val  Leu  Pro  Cys  His  Ser  Ser  Leu  Gln  Pro  Asn  Leu
          25                       30                  35
```

<210> 2
<211> 37
<212> PRT
<213> Paramecium sp.

<400> 2

```
Met  Val  Ser  Met  Phe  Ser  Leu  Ser  Phe  Lys  Trp  Pro  Gly  Phe  Cys  Leu
 1                   5                        10                  15

Phe  Val  Cys  Leu  Phe  Gln  Cys  Pro  Lys  Val  Leu  Pro  Cys  His  Ser  Ser
               20                       25                  30

Leu  Gln  Pro  Asn  Leu
          35
```

<210> 3
<211> 11
<212> PRT
<213> Artificial Sequence

<220>
<223> Designed peptide based on size and polarity to act as a linker between the alpha and beta chains of Protein XYZ.

<400> 3

```
Met  Val  Asn  Leu  Glu  Pro  Met  His  Thr  Glu  Ile
 1                   5                        10
```

<210> 4
<400> 4
000

[Annex VIII follows]